



**Course:** FEN 5201 Data Science for Quantitative Finance  
**Credits:** 3 Credits / Graduate  
**Prereqs:** None  
**Instructor:** Eyal Beigman, PhD

## COURSE OVERVIEW

---

Data science, at its heart, is the art of solving problems and gaining insights using data and models. In this course we will practice solving real world problems using data, applying models and algorithms that are commonly used today. We will practice identifying appropriate sources of data, harvesting the data, ingesting it into a system, and organizing it in a proper data structure. We will practice exploratory data analysis and look at the most important descriptive, predictive, and prescriptive analytic techniques for analytical financial modeling. Throughout this case-based course, there will be an emphasis on asking good guiding questions and effectively communicating results.

## COURSE LEARNING OUTCOMES

---

By the end of this course, students will be able to:

- Identify appropriate sources of data
- Ingest the data into the system using appropriate data structures
- Clean, smooth and transform as appropriate to prepare the data for analysis
- Create appropriate models for different financial applications
- Interpret and effectively communicate results

## REQUIRED TEXTS

---

- Harvard Business School case pack for *Data Science for QE*  
The pack must be purchased and downloaded through the following link:  
<https://hbsp.harvard.edu/import/604441>
- Provost, Foster, Tom Fawcett: *Data Science for Business*.

## ASSIGNMENTS & GRADING

Exercises are to be written in Python in a Jupyter notebooks. They will be published on Canvas and GitHub, The exercise template notebook should be downloaded from GitHub and the final solution submitted via GitHub. This will give students a publicly available portfolio of increasingly complex projects to help showcase their skills to employers.

**Evaluation Criteria.** All course projects will be evaluated like work assignments from a demanding employer. The primary evaluation basis is adherence to the deliverables stated in each assignment's functional requirements. To achieve a top grade, students must also adhere to best practices for:

- Software engineering principles: reproducibility, following [appropriate coding guidelines](#), [DRY](#)



- Business criteria (clearly written, concise, and relevant supporting text, logical flow, spelling, grammar; presentations at appropriate level of detail for audience)

Assignment	Grading
<b>Discussion</b> The weekly discussions will focus on topics covered in the course, implementation, applications, and cases in the pack  <i><b>You are expected to read the cases ahead of class and be prepared for a discussion</b></i>	20%
<b>Exercises (4x8)</b> <ul style="list-style-type: none"> <li>• Exercise #1: Deflategate - Collecting data from the web</li> <li>• Exercise #2: Lending Club A - Cleaning and organizing data</li> <li>• Exercise #3. Lending Club B - Classifiers</li> <li>• Exercise #4. Pump and Dump –</li> </ul>	32%
<b>Mid Term Exam</b>	23%
<b>Final Project (18) and Presentation (12)</b> Working individually or in pairs, students will identify an interesting economic problem, compile a relevant dataset, perform an analysis of the problem using one or more of the methods covered in the course to gain insight into the problem, and offer a solution. Students will present their final projects at a “business level” to their peers.	25%

#### GRADING SCALE

Quality of Performance	Letter Grade	Range %	GPA/ Quality Pts.
Excellent - work is of exceptional quality	A	93 - 100	4
	A-	90 - 92.9	3.7
Good - work is above average	B+	87 - 89.9	3.3
Satisfactory	B	83 - 86.9	3
Below Average	B-	80 - 82.9	2.7
Poor	C+	77 - 79.9	2.3
	C	70 - 76.9	2
Failure	F	< 70	0



## COURSE SCHEDULE

Students should expect to spend at a minimum 10 hours each week on this course.

WEEK	TOPICS	SCHEDULE OF READINGS & MAJOR ASSIGNMENTS
Week 1	Introduction and course overview	Readings: <ul style="list-style-type: none"> <li>HBR - The Sexiest job in the 21st Century</li> <li>Provost, Fawcett, Chapters 1,2</li> </ul>
Week 2	Data collection <ul style="list-style-type: none"> <li>scrapping RESTful, HTML and XML</li> <li>API connectivity - JSON</li> </ul>	Readings: <ul style="list-style-type: none"> <li>HBS - Deflategate</li> <li>VanderPlas - A Whirlwind Tour of Python</li> </ul>
Week 3	Data structure <ul style="list-style-type: none"> <li>Numpy, Pandas</li> </ul>	Readings: <ul style="list-style-type: none"> <li>HBS - Lending Club (A)</li> <li>McKinney, Chapters 1- 6, 8</li> </ul> <p>Exercise 1 due</p>
Week 4	Data cleaning <ul style="list-style-type: none"> <li>filling missing data</li> <li>data consistency</li> <li>outliers</li> </ul>	Readings: <ul style="list-style-type: none"> <li>McKinney, Chapters 7, 9</li> </ul>
Week 5	Predictive models, classification	Readings: <ul style="list-style-type: none"> <li>HBS - Lending Club (B)</li> <li>Provost, Fawcett, Chapters 3</li> </ul> <p>Exercise 2 due</p>

Week 6	Model fitting, regression	Readings: <ul style="list-style-type: none"> <li>Provost, Fawcett, Chapters 4</li> </ul>
Week 7	Overfitting and avoidance	Readings: <ul style="list-style-type: none"> <li>Provost, Fawcett, Chapters 5</li> </ul> <p>Exercise 3 due</p>
Week 8	Model Evaluation	Readings: <ul style="list-style-type: none"> <li>Provost, Fawcett, Chapters 7,8</li> </ul>
Week 9	<b>Midterm</b> + Unsupervised methods, k-nearest neighbors	Readings: <ul style="list-style-type: none"> <li>Provost, Fawcett, Chapters 6</li> </ul> <p>Initial draft of project proposal due</p>
Week 10	Financial Data <ul style="list-style-type: none"> <li>Candelstick</li> <li>Tick Data</li> <li>Equity / FX / Crypto</li> </ul>	Readings: <ul style="list-style-type: none"> <li>HBS - Note on Blockchain and Bitcoin</li> <li>McKinney, Chapters 10, 11</li> </ul>
Week 11	Bayesian Methods	Readings: <ul style="list-style-type: none"> <li>Provost, Fawcett, Chapters 9</li> </ul>



		Final draft of project proposal due
Week 12	Text Analysis	Readings <ul style="list-style-type: none"><li>Provost, Fawcett, Chapters 10</li></ul>
Week 13	Data Science and Business Strategy	Readings Provost, Fawcett, Chapters 13 Exercise 4 due
Week 14	Course Recap	
Week 15	FINAL PROJECTS AND PRESENTATIONS DUE	

#### UNIVERSITY POLICIES & RESOURCES

---

##### **ACCESSIBILITY AND ACCOMODATIONS**

The Office of Disability Services collaborates with students, faculty and staff to provide reasonable accommodations and services to students with disabilities. Students with disabilities who are enrolled in this course and who will be requesting documented disability-related accommodations should make an appointment with the Office of Disability Services, (646) 592-4132, rkohn1@yu.edu, during the first week of class. Once you have been approved for accommodations, please submit your accommodation letter to ensure the successful implementation of those accommodations. For more information, please visit:

<http://yu.edu/Student-Life/Resources-and-Services/Disability-Services/>

##### **ACADEMIC INTEGRITY**

The submission by a student of any examination, course assignment, or degree requirement is assumed to guarantee that the thoughts and expressions therein not expressly credited to another are literally the student's own. Evidence to the contrary will result in appropriate penalties. For more information, visit [http://yu.edu/registrar/grad-](http://yu.edu/registrar/grad-catalog/)

[catalog/http://yu.edu/registrar/grad-catalog/](http://yu.edu/registrar/grad-catalog/)

<http://yu.edu/registrar/grad-catalog/>

##### **STUDENT SUPPORT SERVICES**

If you need any additional help, please visit Student Support Services:

<http://yu.edu/academics/services/><http://yu.edu/academics/services/>